

Effects of parameter estimation on prediction densities: a bootstrap approach

Lorenzo Pascual, Juan Romo, Esther Ruiz*

Departamento de Estadística y Econometría, Universidad Carlos III de Madrid, C/Madrid, 126, 28903 Getafe, Madrid, Spain

Abstract

We use a bootstrap procedure to study the impact of parameter estimation on prediction densities, focusing on seasonal ARIMA processes with possibly non normal innovations. We compare prediction densities obtained using the Box and Jenkins approach with bootstrap densities which may be constructed either taking into account parameter estimation variability or using parameter estimates as if they were known parameters. By means of Monte Carlo experiments, we show that the average coverage of the intervals is closer to the nominal value when intervals are constructed incorporating parameter uncertainty. The effects of parameter estimation are particularly important for small sample sizes and when the error distribution is not Gaussian. We also analyze the effect of the estimation method on the shape of prediction densities comparing prediction densities constructed when the parameters are estimated by Ordinary Least Squares (OLS) and by Least Absolute Deviations (LAD). We show how, when the error distribution is not Gaussian, the average coverage and length of intervals based on LAD estimates are closer to nominal values than those based on OLS estimates. Finally, the performance of the bootstrap intervals is illustrated with two empirical examples. © 2001 International Institute of Forecasters. Published by Elsevier Science B.V. All rights reserved.

Keywords: Forecasting; Least absolute deviations; Non normal distributions; Ordinary least squares

1. Introduction

Our main goal in this paper is to study the impact of parameter estimation on prediction densities and we use the bootstrap as a device to assess its relevance. The interest of building prediction intervals which are able to incorporate the uncertainty due to parameter estimation

has been stressed by Chatfield (1993). In the standard approach to construct prediction intervals, based on Box and Jenkins (1976), prediction errors are assumed to be Gaussian and intervals are obtained with center at the point linear predictor and conditioning on parameter estimates. Consequently, Box and Jenkins (BJ) intervals do not take into account the variability due to parameter estimation and may have coverage which is different from the nominal one when the errors are not Gaussian. Alternatively, prediction intervals can be built using

*Corresponding author. Tel.: +34-916-249-851; fax: +34-916-249-849.

E-mail address: ortega@est-econ.uc3m.es (E. Ruiz).

bootstrap procedures. Bootstrap intervals can incorporate the variability due to parameter estimation without assuming any particular distribution for the errors. We analyze the effect of parameter estimation on the shape of prediction densities using the bootstrap procedure proposed by Pascual, Romo and Ruiz (1998) for $ARIMA(p, d, q)$ models.

First, estimating the parameters by conditional Quasi-Maximum Likelihood (QML), we compare bootstrap intervals (PRR) constructed taking into account parameter variability with intervals obtained by using parameter estimates as if they were the true parameters. The latter approach will be referred to as conditional bootstrap (CB). We compare average coverage and length of BJ, CB and PRR intervals. The difference between BJ and CB intervals could be assignable to the deviation of the innovation distribution from the Gaussian assumption. The difference between CB and PRR intervals could be due to parameter estimation uncertainty. Consequently, we can distinguish between the two sources which could affect the precision of prediction intervals when the model is known. As expected, given that the conditional QML estimator is consistent, the variability due to parameter estimation should be taken into account in the construction of prediction intervals when the series sample size is not large enough. In this case, intervals obtained by conditioning on parameter estimates have average coverage lower than the nominal value. As the sample size increases, the effect of parameter estimation is less important.

We also study the effect of the estimation method on the shape of prediction densities. In particular, we consider the prediction of future values of $ARI(p, d)$ processes and compare prediction intervals obtained when estimating by Ordinary Least Squares (OLS) and by Least Absolute Deviations (LAD). When the error distribution is not Gaussian, prediction densities based on the LAD estimator have, in general,

shapes closer to the corresponding empirical prediction densities. As a second goal of this paper, we show how the bootstrap procedure proposed by Pascual et al. (1998) can be extended to construct prediction intervals in multiplicative seasonal ARIMA models.

The paper is organized as follows. First, in Section 2 we describe the bootstrap procedure proposed by Pascual et al. (1998) to construct prediction intervals. Then, Section 3 contains the Monte Carlo results on the effects of parameter variability on the shape of prediction densities when seasonal ARIMA models are estimated by conditional QML by maximizing the conditional Gaussian likelihood function. Also, we carry out experiments to assess the effects of the method used to estimate the model parameters on prediction intervals. In Section 4, we apply the bootstrap PRR procedure to obtain prediction densities for two real time series: monthly observations of the Italian Industrial Production Index and levels of a luteinizing hormone measured on a healthy woman. Finally, Section 5 contains the conclusions and some suggestions for further research.

2. Bootstrap prediction intervals

We now describe the bootstrap procedure proposed in Pascual et al. (1998) to construct prediction intervals for future values of series generated by $ARIMA(p, d, q)$ processes given by

$$\nabla^d y_t = \phi_0 + \phi_1 \nabla^d y_{t-1} + \cdots + \phi_p \nabla^d y_{t-p} + a_t + \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q}, \quad (1)$$

where a_t is a white noise process, ∇ is the difference operator such that $\nabla y_t = y_t - y_{t-1}$ and $(\phi_0, \phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q)$ are unknown parameters. From an observed series $\{y_1, y_2, \dots, y_T\}$, the parameters can be estimated by a consistent estimator, for example conditional

QML. Given $(\hat{\phi}_0, \hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\theta}_1, \dots, \hat{\theta}_q)$, the residuals are calculated by the following recursion

$$\begin{aligned} \hat{a}_t = & \nabla^d y_t - \hat{\phi}_0 - \hat{\phi}_1 \nabla^d y_{t-1} - \dots \\ & - \hat{\phi}_p \nabla^d y_{t-p} - \hat{\theta}_1 \hat{a}_{t-1} - \dots - \hat{\theta}_q \hat{a}_{t-q}, \\ & t = p + d + 1, \dots, T, \end{aligned} \quad (2)$$

where the residuals corresponding to periods of time $t = 0, -1, -2, \dots$ are set equal to zero. Denote by \hat{F}_a the empirical distribution function of the centered residuals. Given a set of $p + d$ initial values of the variable y_t , say $\{y_1, \dots, y_{p+d}\}$, a bootstrap replicate of the series $\{y_1^*, \dots, y_T^*\}$ is constructed by the following equation

$$\begin{aligned} \nabla^d y_t^* = & \hat{\phi}_0 + \sum_{j=1}^p \hat{\phi}_j \nabla^d y_{t-j}^* + \sum_{j=1}^q \hat{\theta}_j \hat{a}_{t-j}^* + \hat{a}_t^*, \\ & t = p + d + 1, \dots, T, \end{aligned} \quad (3)$$

where $y_t^* = y_t$, $t = 1, \dots, p + d$ and $\hat{a}_{1+p+d-q}^*, \dots, \hat{a}_T^*$ are random draws from \hat{F}_a . Once the parameters of this bootstrap series are estimated, say $(\hat{\phi}_0^*, \hat{\phi}_1^*, \dots, \hat{\phi}_p^*, \hat{\theta}_1^*, \dots, \hat{\theta}_q^*)$, the bootstrap forecast k steps ahead is obtained as follows,

$$\begin{aligned} \nabla^d y_{T+k}^* = & \hat{\phi}_0^* + \sum_{j=1}^p \hat{\phi}_j^* \nabla^d y_{T+k-j}^* \\ & + \sum_{j=1}^q \hat{\theta}_j^* \hat{a}_{T+k-j}^* + \hat{a}_{T+k}^*, \\ & k = 1, 2, \dots \end{aligned} \quad (4)$$

where $y_{T+k-j}^* = y_{T+k-j}$, $j \geq k$, and $\hat{a}_{T+k-j}^* = \hat{a}_{T+k-j}$, $j \geq k$, i.e., the last $p + d$ observations of the series and the last q residuals are fixed in order to obtain the prediction density conditional on the observed data. Finally, in expression (4), \hat{a}_{T+k-j}^* , $j < k$ are random draws from \hat{F}_a . As an illustration, we consider an ARIMA(1,1,1) model without constant term

$$\nabla y_t = \phi \nabla y_{t-1} + a_t + \theta a_{t-1}. \quad (5)$$

Once the parameters of model (5) have been estimated and the bootstrap draws $\hat{a}_2^*, \dots, \hat{a}_T^*$ are available, a bootstrap replicate of the series is constructed by

$$\begin{aligned} y_t^* = & (1 + \hat{\phi}) y_{t-1}^* - \hat{\phi} y_{t-2}^* + \hat{a}_t^* + \hat{\theta} \hat{a}_{t-1}^*, \\ & t = 3, \dots, T, \end{aligned} \quad (6)$$

where $y_1^* = y_1$ and $y_2^* = y_2$. Then, bootstrap estimates $\hat{\phi}^*$ and $\hat{\theta}^*$ are obtained for the bootstrap series and bootstrap replicates of future values of the series are generated by

$$\begin{aligned} y_{T+1}^* = & (1 + \hat{\phi}^*) y_T^* - \hat{\phi}^* y_{T-1}^* \\ & + \hat{a}_{T+1}^* + \hat{\theta}^* \hat{a}_T^*, \\ y_{T+2}^* = & (1 + \hat{\phi}^*) y_{T+1}^* - \hat{\phi}^* y_T^* \\ & + \hat{a}_{T+2}^* + \hat{\theta}^* \hat{a}_{T+1}^*, \\ y_{T+3}^* = & (1 + \hat{\phi}^*) y_{T+2}^* - \hat{\phi}^* y_{T+1}^* \\ & + \hat{a}_{T+3}^* + \hat{\theta}^* \hat{a}_{T+2}^*, \dots \end{aligned} \quad (7)$$

Notice that in the recursions above \hat{a}_T is kept fixed in the different bootstrap replicates of y_{T+1}^* while \hat{a}_{T+1}^* changes from one to another replicate.

This procedure is repeated B times to obtain a set of B bootstrap replicates $\{y_{T+k}^{*(1)}, \dots, y_{T+k}^{*(B)}\}$. Finally, the prediction limits are defined as the quantiles of the bootstrap distribution function of y_{T+k}^* , i.e., if $G^*(h) = \Pr(y_{T+k}^* \leq h)$ is the distribution function of y_{T+k}^* and its Monte Carlo estimate is $G_B^*(h) = \#(y_{T+k}^{*(b)} \leq h)/B$, a $100\alpha\%$ prediction interval for Y_{T+k}^* is given by

$$[L_B^*, U_B^*] = \left[Q_B^* \left(\frac{1-\alpha}{2} \right), Q_B^* \left(\frac{1+\alpha}{2} \right) \right] \quad (8)$$

where $Q_B^* = G_B^{*-1}$.

Notice that in the procedure just described, the last $p + d$ observations of the series and the final q residuals are fixed in all bootstrap replicates of future values so we can obtain the prediction density conditional on the observed sample. However, with the exception of the first

$p + d$ initial values, we do not fit any observation when generating bootstrap replicates of the series used to obtain bootstrap estimates of the parameters of the model.

In the bootstrap procedure proposed by Thombs and Schucany (1990) for AR(p) processes, they fit the last p observed values of y_t to obtain bootstrap replicates of the series and, consequently, they need to use the backward representation of the AR(p) process. In the backward representation, the process y_t is expressed as a linear combination of future values plus an error term. For example, for an AR(p) process the backward representation is given by

$$y_t = \phi_0 + \phi_1 y_{t+1} + \cdots + \phi_p y_{t+p} + e_t. \quad (9)$$

The errors of the backward representation can be estimated by

$$\begin{aligned} \hat{e}_t &= y_t - \hat{\phi}_0 - \hat{\phi}_1 y_{t+1} - \cdots - \hat{\phi}_p y_{t+p}, \\ t &= 1, \dots, T-p. \end{aligned} \quad (10)$$

Then, bootstrap replicates of the series, $\{y_1^*, \dots, y_T^*\}$, generated to obtain bootstrap estimates of the parameters can be generated with the same last p values by choosing $y_{T-j}^* = y_{T-j}$, $j = 0, 1, \dots, p-1$ and generating the remainder values of y_{T-j}^* by the following recursion

$$\begin{aligned} y_t^* &= \hat{\phi}_0 + \sum_{j=1}^p \hat{\phi}_j y_{t+j}^* + \hat{e}_t^*, \\ t &= T-p, \dots, 1, \end{aligned} \quad (11)$$

where \hat{e}_t^* are random draws from \hat{F}_e , the empirical distribution of the centered and rescaled backward residuals; see Thombs and Schucany (1990) and for the rescaling see Stine (1987). One problem of the backward representation when the forward errors are non-Gaussian, is that even if they are independent, the backward residuals are not independent, merely uncorrelated. Furthermore, the need to use the backward representation restricts the use of the

Thombs and Schucany (1990) procedure to models with such representation excluding, for example, generalized autoregressive conditional heteroscedasticity (GARCH) models. Therefore, the main advantage of the method just described over the technique in Thombs and Schucany (1990) is that the computational burden associated with resampling through the backward representation is avoided. Consequently, the PRR bootstrap procedure can be easily applied to models with moving average components while the procedure proposed by Thombs and Schucany (1990) can only be directly applied to autoregressive models. In Pascual et al. (1998) can be seen a proof of the asymptotic validity of this bootstrap resampling and a Monte Carlo comparison between both proposals.

Alternatively, the bootstrap procedure just described could be also applied to construct prediction intervals conditional on the parameter estimates (CB). In this case, the parameters are estimated once and these estimates are used in the calculation of all bootstrap forecasts $\nabla^d y_{T+k}^*$. Therefore, it is not necessary to generate bootstrap replicates of the series as in (3) and the bootstrap forecast k steps ahead depends only on the resampled residuals and is given by

$$\begin{aligned} \nabla^d y_{T+k}^* &= \hat{\phi}_0 + \sum_{j=1}^p \hat{\phi}_j \nabla^d y_{T+k-j}^* \\ &\quad + \sum_{j=1}^q \hat{\theta}_j \hat{a}_{T+k-j}^* + \hat{a}_{T+k}^*, \\ k &= 1, 2, \dots, \end{aligned} \quad (12)$$

where y_{T+k-j}^* and \hat{a}_{T+k-j}^* are defined as in (4). Since the parameter estimates are kept fixed in all bootstrap replicates of future values, the CB prediction intervals do not incorporate the uncertainty due to parameter estimation. In the case of AR(p) processes, the conditional bootstrap was proposed by Cao et al. (1997).

Finally, it is important to mention that for the bootstrap procedures previously described, PRR

and CB, there are alternative methods to build prediction intervals to the one proposed in (8). In particular, Hall (1992) describes, among others, two alternative intervals. The first one, called the percentile interval, is given by

$$\left[2\hat{y}_{T+k} - t_{1-\frac{\alpha}{2}}^*, 2\hat{y}_{T+k} - t_{\frac{\alpha}{2}}^* \right] \quad (13)$$

where \hat{y}_{T+k} is the point linear predictor of y_{T+k} and t^* is a percentile of $G^*(h)$. Secondly, prediction intervals can be built by the percentile- t method as follows

$$\left[\hat{y}_{T+k} - \hat{\sigma}_{T+k} t_{1-\frac{\alpha}{2}}^{\dagger}, \hat{y}_{T+k} - \hat{\sigma}_{T+k} t_{\frac{\alpha}{2}}^{\dagger} \right] \quad (14)$$

where $\hat{\sigma}_{T+k}$ is the usual estimate of the standard deviation of the k -steps ahead prediction errors based on the estimated parameters $(\hat{\phi}_0, \hat{\phi}_1, \dots, \hat{\phi}_p, \hat{\theta}_1, \dots, \hat{\theta}_q)$ and the residual standard deviation and t^{\dagger} is a percentile of the distribution of $(y_{T+k}^* - \hat{y}_{T+k})/\hat{\sigma}_{T+k}^*$ where $\hat{\sigma}_{T+k}^*$ is the bootstrap counterpart of $\hat{\sigma}_{T+k}$. Hall (1992) points out that intervals (13) and (14) are not able to deal with skewness of the prediction error distribution and can be improved by appropriate transformations. However, he stresses that the interval in (8) is transformation-respecting and consequently, is not affected by asymmetry of the error distribution.

3. Effects of estimation on prediction densities

In this section, several Monte Carlo experiments are carried out to study the effect of parameter estimation variability on the shape of estimated prediction densities. Prediction densities are constructed by the bootstrap procedure described in the previous section, either conditioning on parameter estimates (CB) or introducing the variability due to parameter estimation (PRR). The focus is on prediction of future values of multiplicative seasonal ARIMA(p, d, q) \times (P, D, Q)_s processes, where s is the season-

al period. First, we will consider a stationary ARMA(1,1) process. Then, prediction densities will be constructed for a model with unit roots. Finally, we will analyze the properties of both bootstrap procedures in models with seasonal components. For each model considered, we generate artificial series with several choices of error distributions, in particular, Gaussian, Student-t with 3 degrees of freedom, Chi-squared with four degrees of freedom and exponential errors. In all cases, we have centered and rescaled the errors to have zero mean and unit variance. All the models are estimated by conditional QML that coincides with OLS when the model lacks of a moving average component. All computations have been carried out in a HP-UX C360 workstation, using Fortran 77 and the corresponding subroutines of Numerical Recipes by Press, Flannery, Teukolsky and Vetterling (1986). In particular, Gaussian, Student-t and Chi-squared errors are generated using the subroutine “*gasdev*” and the corresponding transformations in each case. Exponential errors are generated using uniform random numbers generated by subroutine “*ran2*” and transforming them as appropriate. The numerical optimization of the Gaussian log-likelihood function has been carried out using the subroutine “*amoeba*” with the maximum allowed function evaluations set equal to 5000 and the fractional convergence tolerance set equal to 10^{-6} . Finally, the subroutine used to obtain the LAD estimates of the parameters is “*medfi*” with the convergence tolerance set equal to 10^{-7} .

3.1. ARMA processes

To illustrate the effect of parameter variability on estimated prediction densities of stationary ARMA processes, we generated 1000 time series with the following ARMA(1,1) process

$$y_t = 0.7y_{t-1} + a_t - 0.3a_{t-1}, \quad (15)$$

where a_t is Gaussian. For each series, we compute the empirical prediction density by generating future values of y_{T+k} conditional on $\{y_1, y_2, \dots, y_T\}$. We also calculate the bootstrap prediction densities obtained conditioning on the parameter estimates (CB) and by using the PRR technique described in the previous section. Both densities are based on 999 bootstrap replicates. Finally, we constructed prediction intervals based on the Box and Jenkins procedure (BJ). Notice that the difference between BJ and CB intervals could be associated with departures of the innovation distribution from Gaussianity. On the other hand, the differences between CB and PRR intervals are assignable to the uncertainty in the estimation of the parameters. The average coverage, the average coverage for each tail and the average length of intervals constructed with a 95% nominal cover-

age are reported in Table 1 for predictions one and three steps ahead and $T = 25, 50$ and 100 . For the Gaussian innovation distribution in this table, CB intervals have lower average coverage than PRR intervals, the latter having average coverage closer to the nominal value. Note that the average length of CB intervals is also shorter than the empirical length. This effect is more evident for small sample sizes. Consequently, it seems that for relatively small sample sizes, it is important to include the uncertainty due to parameter estimation in prediction intervals in order to obtain coverages closer to the nominal values. As expected, since the conditional QML estimator is consistent, CB and PRR intervals get closer in terms of coverage and length as the sample size increases. The results are similar for predictions made one and three steps ahead. Comparing BJ and PRR

Table 1
Monte Carlo results for model $y_t = 0.7y_{t-1} + a_t - 0.3a_{t-1}$ with Gaussian errors^a

Lead time	Sample Size	Method	Average coverage(se)	Coverage below/above	Average length
1	n	Empirical	95%	2.5%/2.5%	3.92
	25	BJ	92.74(0.05)	3.8/3.4	3.90(0.57)
		CB	90.68(0.07)	4.87/4.44	3.84(0.70)
		PRR	92.70(0.05)	3.7/3.6	3.99(0.70)
	50	BJ	94.04(0.03)	3.1/2.8	3.92(0.39)
		CB	92.12(0.04)	4.04/3.84	3.79(0.52)
		PRR	93.46(0.03)	3.3/3.2	3.93(0.51)
	100	BJ	94.49(0.02)	2.7/2.7	3.92(0.29)
		CB	93.66(0.03)	3.06/3.28	3.87(0.39)
		PRR	94.04(0.02)	2.91/3.05	3.91(0.37)
3	n	Empirical	95%	2.5%/2.5%	4.35
	25	BJ	93.59(0.05)	3.4/3.0	4.48(0.79)
		CB	90.79(0.07)	4.81/4.40	4.18(0.74)
		PRR	93.14(0.04)	3.4/3.4	4.38(0.77)
	50	BJ	94.39(0.03)	2.9/2.8	4.44(0.54)
		CB	92.97(0.04)	3.67/3.36	4.26(0.55)
		PRR	93.84(0.03)	3.1/3.1	4.36(0.55)
	100	BJ	94.78(0.02)	2.6/2.6	4.42(0.39)
		CB	93.95(0.25)	2.93/3.12	4.30(0.40)
		PRR	94.27(0.02)	2.8/2.9	4.34(0.41)

^a Quantities in parentheses are standard deviations.

intervals, it is surprising to observe that, under Gaussian innovations, although BJ intervals do not incorporate estimation uncertainty, their properties are similar to the PRR intervals even for small sample sizes.

Table 2 reports the Monte Carlo results for 80% prediction intervals for series generated by model (15) with innovations generated by a Student-t distribution with 3 degrees of freedom rescaled to have unit variance. It can be seen that prediction intervals are improved when they include the variability due to parameter estimation. Differences between CB and PRR intervals are larger than for Gaussian errors. As expected, these differences are smaller, the bigger the sample size. Table 2 also includes the results of BJ intervals. The average coverage of BJ intervals is over nominal values and the average length is also over the empirical length. Notice,

that this distortion does not disappear when the sample size increases. The results are similar for one and three steps ahead predictions. Consequently, as pointed out by Pascual et al. (1998), BJ intervals are clearly distorted when the error distribution is not Gaussian.

Table 3 reports the Monte Carlo results for the same model but with innovations generated by an exponential distribution centered to have zero mean. Once more, differences between CB and PRR intervals are larger than for Gaussian innovations. Therefore, when the innovations are not normal and estimation is carried out by conditional QML, it seems important to include the variability due to parameter estimation in prediction intervals. For the sake of comparison, Table 3 also includes the results of BJ intervals which are not able to deal with the asymmetry of the error distribution. As an illustration, Fig.

Table 2
Monte Carlo results for model $y_t = 0.7y_{t-1} + a_t - 0.3a_{t-1}$ with Student-t errors^a

Lead time	Sample size	Method	Average coverage(se)	Coverage below/above	Average length
1	n	Empirical	80%	10%/10%	1.90
	25	BJ	82.04(0.10)	8.70/9.26	2.43(1.19)
		CB	74.14(0.12)	12.82/13.04	1.91(0.51)
		PRR	78.78(0.10)	10.31/10.91	2.09(0.57)
	50	BJ	84.37(0.07)	7.78/7.85	2.42(0.79)
		CB	77.21(0.08)	11.30/11.49	1.92(0.35)
		PRR	79.51(0.07)	10.15/10.34	2.00(0.35)
	100	BJ	86.11(0.05)	6.98/6.90	2.47(0.67)
		CB	78.77(0.05)	10.64/10.59	1.91(0.25)
		PRR	79.60(0.04)	10.19/10.21	1.94(0.24)
3	n	Empirical	80%	10%/10%	2.25
	25	BJ	81.67(0.10)	9.22/9.11	2.79(1.36)
		CB	75.69(0.11)	12.24/12.07	2.32(0.81)
		PRR	79.20(0.09)	10.36/10.44	2.44(0.78)
	50	BJ	83.42(0.07)	8.31/8.27	2.74(0.89)
		CB	77.51(0.08)	11.16/11.34	2.27(0.42)
		PRR	79.12(0.06)	10.39/10.49	2.32(0.43)
	100	BJ	85.00(0.05)	7.62/7.38	2.79(0.77)
		CB	78.68(0.05)	10.75/10.57	2.25(0.30)
		PRR	79.34(0.05)	10.39/10.27	2.28(0.31)

^a Quantities in parenthesis are standard deviations.

Table 3

Monte Carlo results for model $y_t = 0.7y_{t-1} + a_t - 0.3a_{t-1}$ with exponential errors^a

Lead time	Sample size	Method	Average coverage(se)	Coverage below/above	Average length
1	n	Empirical	95%	2.5%/2.5%	3.64
	25	BJ	92.97(0.06)	0.58/6.44	3.83(1.02)
		CB	89.52(0.12)	6.34/4.14	3.86(1.30)
		PRR	93.28(0.08)	2.8/3.9	4.05(1.39)
	50	BJ	94.09(0.03)	0.01/5.81	3.88(0.79)
		CB	90.92(0.09)	5.22/3.86	3.64(0.96)
		PRR	94.27(0.06)	2.2/3.5	3.86(1.04)
	100	BJ	94.44(0.02)	0.0/5.56	3.89(0.55)
		CB	93.14(0.06)	3.64/3.21	3.65(0.64)
		PRR	94.91(0.05)	1.97/3.12	3.74(0.66)
3	n	Empirical	95%	2.5%/2.5%	4.20
	25	BJ	93.64(0.04)	0.39/5.96	4.40(1.25)
		CB	89.40(0.10)	6.0/4.59	4.14(1.29)
		PRR	93.25(0.06)	2.7/4.07	4.39(1.39)
	50	BJ	94.28(0.03)	0.02/5.53	4.38(0.91)
		CB	91.32(0.07)	5.0/3.68	4.18(1.00)
		PRR	93.48(0.05)	3.08/3.4	4.30(1.01)
	100	BJ	94.83(0.02)	0.003/5.14	4.38(0.64)
		CB	93.06(0.05)	3.81/3.12	4.17(0.67)
		PRR	93.94(0.04)	2.98/3.08	4.22(0.69)

^aQuantities in parenthesis are standard deviations.

1 represents the empirical, BJ, CB and PRR densities obtained for one step ahead predictions of one of the series generated by model (15) with exponential innovations and $T = 100$. This figure shows that the density constructed by taking into account the variability due to parameter estimation is much closer to the empirical density than when parameter estimates are considered as fixed. Furthermore, it can be seen that BJ density is clearly distorted.

Finally, Table 4 reports the results obtained for one-step ahead 95% prediction intervals for model (15) with Gaussian innovations, built using the PRR procedure by the expression in (8) and by the two alternative methods described by Hall (1992), the percentile interval in (13) and the percentile- t interval in (14). The results in Table 4 show that the prediction intervals built by (13) and the intervals in (8)

have very similar properties when the innovations distribution is symmetric. However, if the distribution is Chi-squared or exponential, the intervals built by (13) are not able to capture the asymmetry in the prediction errors. The percentile- t intervals have the same problem for non-symmetric distributions although when the distribution is symmetric and for small sample sizes, the average coverage can be slightly closer to nominal values. However, notice that the percentile- t intervals have average length larger than for intervals built by (8) which are closer to nominal. Since both the percentile and percentile- t intervals in (13) and (14) are centered at the linear predictor, it may happen that when the innovation distribution is not symmetric, the intervals are not adequately centered and they might be improved by not centering the interval at a particular value as in (8). In any

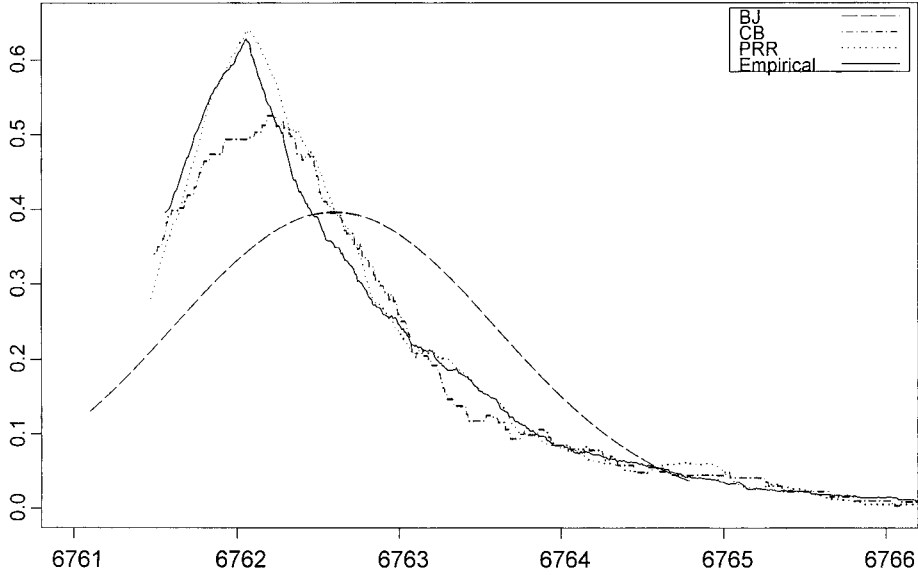


Fig. 1. Densities of one-step ahead predictions of one series of size 100 generated by model $y_t = 0.7y_{t-1} + a_t - 0.3a_{t-1}$ with exponential innovations.

case, the differences in coverage between the intervals in (8) and the percentile- t intervals only appear for symmetric distributions and very small sample sizes ($T=25$). When the sample size is moderate ($T=100$), both intervals have similar coverage for symmetric distributions with the intervals in (8) having shorter length. On the other hand, the intervals in (8) are clearly superior for asymmetric distributions for any sample size. On top of that, the computations needed to construct the intervals in (8) are simpler than for the percentile- t intervals. Therefore, we recommend to use in practice the intervals in (8) unless the sample size is very small and there is evidence that the error distribution is symmetric. Consequently, we will obtain prediction intervals by expression (8) for all the simulations and the empirical application.

3.2. Unit root processes

To analyze how the presence of unit roots may affect the previous conclusions, we generated 1000 series from the ARI(2,1) process

$$\nabla^2 y_t = 0.5\nabla^2 y_{t-1} + a_t \quad (16)$$

with exponential innovations. The results for 95% prediction intervals are reported in Table 5 where it can be observed that the one-step ahead intervals have similar behavior to that previously commented, i.e. BJ intervals are not able to capture the asymmetry present in the data and the average coverage of the intervals built conditional on parameter estimates is generally under nominal coverage. Also, notice that CB intervals are not able to correctly capture the asymmetry of the error distribution. Finally, PRR intervals have average coverage close to the nominal value and they capture properly the error prediction asymmetry. Notice that when predictions are made three steps ahead, the average coverage of BJ intervals is over the nominal value, implying more uncertainty about the future than they should. Table 6 reports the results obtained when the nominal coverage is 80%. In this case, the behavior of BJ intervals is even worse than before. For example, the average coverage of BJ intervals for three steps

Table 4

Monte Carlo results for model $y_t = 0.7y_{t-1} + a_t - 0.3a_{t-1}$ using PRR and lead time 1. Perc₁ is based in (8), Perc₂ in (13) and Perc- t in (14)^a

Error distribution	Sample size	Bootstrap method	Average coverage(se)	Coverage below/above	Average length
Gaussian	n	Empirical	95%	2.5%/2.5%	3.92
	25	Perc ₁	92.70(0.05)	3.7/3.6	3.99(0.70)
		Perc ₂	91.72(0.07)	4.46/3.38	3.99(0.70)
		Perc- t	94.01(0.06)	3.25/2.74	4.40(0.78)
	100	Perc ₁	94.04(0.02)	2.91/3.05	3.91(0.37)
		Perc ₂	94.07(0.02)	3.0/2.93	3.91(0.37)
		Perc- t	94.73(0.02)	2.67/2.60	4.01(0.37)
Student-t	n	Empirical	95%	2.5%/2.5%	3.68
	25	Perc ₁	92.96(0.06)	3.40/3.64	4.38(2.50)
		Perc ₂	92.59(0.06)	3.52/3.89	4.38(2.50)
		Perc- t	94.33(0.05)	2.71/2.96	4.87(3.44)
	100	Perc ₁	94.25(0.02)	2.79/2.96	3.80/(0.76)
		Perc ₂	94.29(0.02)	2.96/2.75	3.81(0.76)
		Perc- t	95.11(0.02)	2.53/2.36	4.07(0.73)
Exponential	n	Empirical	95%	2.5%/2.5%	3.64
	25	Perc ₁	93.28(0.08)	2.8/3.9	4.05(1.39)
		Perc ₂	88.47(0.06)	0.32/11.21	4.05(1.39)
		Perc- t	91.44(0.05)	0.15/8.41	4.62(1.64)
	100	Perc ₁	94.91(0.05)	1.97/3.12	3.74(0.66)
		Perc ₂	87.42(0.02)	0.00/12.58	3.74(0.66)
		Perc- t	88.57(0.02)	0.00/11.43	3.94(0.67)
Chi-squared	n	Empirical	95%	2.5%/2.5%	3.75
	25	Perc ₁	93.24(0.06)	2.83/3.93	4.04(1.11)
		Perc ₂	89.57(0.06)	0.56/9.87	4.04(1.12)
		Perc- t	92.10(0.05)	0.32/7.58	4.54(1.28)
	100	Perc ₁	94.33(0.04)	2.61/3.06	3.79(0.53)
		Perc ₂	89.48(0.02)	0.00/10.52	3.79(0.53)
		Perc- t	90.31(0.02)	0.00/9.69	3.95(0.54)

^a Quantities in parenthesis are standard deviations.

ahead predictions constructed with 100 observations is 93.54%, i.e. 13.54% larger than nominal. Of course, BJ intervals are not able to capture the asymmetry in the innovation distribution. The behavior of CB and PRR intervals is similar to that for 95% intervals.

The features of the estimated prediction intervals may also depend on the properties of the

estimation method used. Thus, in this subsection we will also compare intervals for $ARI(p, d)$ processes constructed when the parameters of the model are estimated either by OLS or by LAD. The results in Table 7 are based on 1000 series generated by model (12) with exponential innovations and with model parameters estimated by LAD. The results for the same model

Table 5

Monte Carlo results for model $(1 - B)^2 (1 - 0.5B)y_t = a_t$ with exponential errors^a

Lead time	Sample size	Method	Average coverage(se)	Coverage below/above	Average length
1	n	Empirical	95%	2.5%/2.5%	3.65
	25	BJ	93.33(0.04)	0.14/6.5	3.77(1.04)
		CB	88.01(0.12)	7.6/4.4	3.62(1.23)
		PRR	91.02(0.09)	4.7/4.3	3.72(1.26)
	50	BJ	94.03(0.03)	0.11/5.9	3.84(0.72)
		CB	90.15(0.09)	6.07/3.77	3.60(0.94)
		PRR	92.65(0.07)	3.8/3.5	3.70(0.89)
	100	BJ	94.44(0.02)	0.0/5.56	3.87(0.53)
		CB	91.85(0.07)	5.04/3.11	3.65(0.66)
		PRR	93.45(0.06)	3.5/3.05	3.72(0.68)
3	n	Empirical	95%	2.5%/2.5%	19.05
	25	BJ	96.48(0.03)	0.18/3.34	26.09(7.82)
		CB	88.07(0.10)	6.61/5.31	17.34(5.39)
		PRR	90.46(0.08)	4.4/5.1	18.01(5.72)
	50	BJ	97.41(0.02)	0.1/2.6	26.96(5.37)
		CB	91.60(0.07)	4.79/3.61	18.62(3.99)
		PRR	92.75(0.06)	3.7/3.6	18.85(3.90)
	100	BJ	97.75(0.01)	0.0/2.25	27.35(4.02)
		CB	92.96(0.05)	3.91/3.13	18.75(2.90)
		PRR	93.55(0.04)	3.3/3.13	18.89(2.96)

^a Quantities in parentheses are standard deviations.

estimated by OLS were reported in Table 5. Comparing Tables 5 and 7, we observe that when parameters are estimated by LAD, the average coverage of CB and PRR intervals is closer to the nominal value of 95%. For example, with a sample size of 100 and intervals constructed for one-step ahead predictions, the average coverage when OLS is used is 93.45 and when the parameters are estimated by LAD, the average coverage is 94.79. The same holds for three-steps ahead predictions. The bootstrap intervals based on LAD estimates are closer to nominal intervals due perhaps to the better properties of such estimators under non-symmetric distributions. If the parameter estimator has better properties it is expected that the predictions will also be better. Notice that

the BJ intervals behavior is quite similar when estimating either by OLS or by LAD.

3.3. Seasonal models

The focus in this section is on prediction densities of series generated by multiplicative seasonal $\text{ARIMA}(p, d, q) \times (P, D, Q)_s$ processes, where s is the seasonal period. For example, for monthly data, we consider the model

$$\phi_p(L)\Phi_p(L^{12})\nabla^d\nabla_{12}^D Y_t = \theta_q(L)\Theta_Q(L^{12})a_t, \quad (17)$$

where L is the backshift operator such that $L^k y_t = y_{t-k}$ and $\nabla_{12}^D = (1 - L^{12})^D$, and the autoregressive and moving average polynomials $\phi_p(L) = (1 - \phi_1 L - \dots - \phi_p L^p)$, $\Phi_p(L^{12}) =$

Table 6

Monte Carlo results for model $(1 - B)^2 (1 - 0.5B)y_t = a_t$ with exponential errors^a

Lead time	Sample size	Method	Average coverage(se)	Coverage below / above	Average length
1	n	Empirical	80%	10% / 10%	2.19
	25	BJ	84.60(0.10)	3.70 / 11.7	2.46(0.68)
		CB	73.13(0.15)	14.81 / 12.05	2.11(0.60)
		PRR	76.06(0.10)	12.72 / 11.21	2.23(0.61)
	50	BJ	87.44(0.07)	1.6 / 10.98	2.51(0.47)
		CB	76.68(0.12)	12.30 / 11.02	2.16(0.43)
		PRR	77.96(0.11)	11.25 / 10.97	2.20(0.42)
	100	BJ	88.86(0.03)	0.52 / 10.62	2.53(0.35)
		CB	77.96(0.10)	11.56 / 10.48	2.18(0.31)
		PRR	78.62(0.09)	10.99 / 10.4	2.19(0.31)
3	n	Empirical	80%	10% / 10%	2.25
	25	BJ	81.67(0.10)	9.22 / 9.11	2.79(1.36)
		CB	75.69(0.11)	12.24 / 12.07	2.32(0.81)
		PRR	79.20(0.09)	10.36 / 10.44	2.44(0.78)
	50	BJ	83.42(0.07)	8.31 / 8.27	2.74(0.89)
		CB	77.51(0.08)	11.16 / 11.34	2.27(0.42)
		PRR	79.12(0.06)	10.39 / 10.49	2.32(0.43)
	100	BJ	85.00(0.05)	7.62 / 7.38	2.79(0.77)
		CB	78.68(0.05)	10.75 / 10.57	2.25(0.30)
		PRR	79.34(0.05)	10.39 / 10.27	2.28(0.31)

^a Quantities in parentheses are standard deviations.

$(1 - \Phi_1 L^{12} - \dots - \Phi_p L^{12p})$, $\theta_q(L) = (1 + \theta_1 L + \dots + \theta_q L^q)$, $\Theta_Q(L^{12}) = (1 + \Theta_1 L^{12} + \dots + \Theta_Q L^{12Q})$ have all their roots out of the unit circle to ensure stationarity and invertibility. In particular, we generate 1000 series with the following $\text{ARIMA}(0,1,1) \times (0,1,1)_{12}$ model, usually known as *airline* model:

$$\nabla \nabla^{12} y_t = (1 - 0.33L)(1 - 0.82L^{12})a_t, \quad (18)$$

where a_t is a Gaussian innovation. Table 8 reports the results obtained for sample sizes 120 and 240. For Gaussian errors and sample sizes rather large, the properties of the prediction densities constructed by the three methods considered in this paper are rather similar. The results for different innovation distributions and sample sizes are similar to the ones previously commented for models (15) and (16) and,

consequently, they are not reported. The intent of this Monte Carlo experiment is to show how the PRR procedure can be extended to seasonal models with good results. As an illustration, the empirical, BJ, CB and PRR densities of one-step ahead predictions for a particular series of size 240 generated by model (18) appear in Fig. 2 where it can be seen that all densities are very similar.

4. Real data applications

In this section, we study the implementation of the PRR bootstrap procedure to the prediction of future values of two real time series, monthly observations of the Italian Industrial Production Index (IPI) and observations of the levels of a luteinizing hormone from Efron and

Table 7

Monte Carlo results for model $(1 - B)(1 - B^{12})y_t = (1 - 0.33B)(1 - 0.82B^{12})a_t$ with Gaussian errors^a

Lead time	Sample size	Method	Average coverage(se)	Coverage below/above	Average length
1	n	Empirical	95%	2.5%/2.5%	3.92
	120	BJ	95.95(0.02)	2.06/1.99	4.25(0.31)
		CB	95.14(0.03)	2.44/2.42	4.20(0.42)
		PRR	95.26(0.03)	2.39/2.34	4.23(0.41)
	240	BJ	95.66(0.01)	2.20/2.15	4.10(0.20)
		CB	95.15(0.02)	2.38/2.47	4.07(0.03)
		PRR	95.19(0.02)	2.34/2.46	4.08(0.28)
3	n	Empirical	95%	2.5%/2.5%	5.40
	120	BJ	95.89(0.02)	2.09/2.02	5.87(0.58)
		CB	95.18(0.03)	2.42/2.40	5.77(0.60)
		PRR	95.41(0.03)	2.29/2.31	5.87(0.61)
	240	BJ	95.61(0.02)	2.19/2.20	5.65(0.38)
		CB	95.20(0.02)	2.35/2.45	5.60(0.42)
		PRR	95.27(0.02)	2.29/2.43	5.65(0.43)
12	n	Empirical	95%	2.5%/2.5%	9.54
	120	BJ	95.53(0.03)	2.26/2.20	10.38(1.45)
		CB	94.06(0.04)	2.96/2.97	10.22(1.44)
		PRR	94.55(0.04)	2.70/2.71	10.43(1.45)
	240	BJ	95.42(0.02)	2.26/2.32	9.99(0.95)
		CB	94.59(0.03)	2.63/2.77	9.90(1.01)
		PRR	94.82(0.03)	2.49/2.68	10.01(0.99)
24	n	Empirical	95%	2.5%/2.5%	14.42
	120	BJ	96.35(0.03)	1.85/1.80	16.72(2.61)
		CB	94.59(0.05)	2.71/2.70	16.50(2.64)
		PRR	95.77(0.04)	2.11/2.12	17.50(2.64)
	240	BJ	95.97(0.02)	1.99/2.04	15.64(1.67)
		CB	94.92(0.03)	2.47/2.61	15.52(1.73)
		PRR	95.66(0.020)	2.09/2.25	16.12(1.73)

^a Quantities in parentheses are standard deviations.

Tibshirani (1993), previously analyzed by Diggle (1990). The Italian IPI observed monthly from January 1983 to September 1998 can be seen in Fig. 3 and it presents a strong seasonal component and a stochastic trend. The first 165 observations of the series, corresponding to the period up to September 1996, are used to estimate the $ARIMA(p, d, q) \times (P, D, Q)_{12}$ model which describes the dynamic behavior of the Italian IPI. The last 24 observations are used

to assess the predictive performance of the Box–Jenkins and bootstrap prediction intervals.

Before estimating the model, the effects of several outliers have been removed from the original series using the program TRAMO; see Gómez and Maravall (1996). The model estimated by conditional QML from the series without outliers is given by

$$\nabla \nabla_{12} y_t = (1 - 0.59L)_{(0.07)} (1 - 0.57L^{12})_{(0.07)} a_t. \quad (19)$$

Table 8

Monte Carlo results for model $(1 - B)^2 (1 - 0.5B)y_t = a_t$ with exponential errors and LAD estimation^a

Lead time	Sample size	Method	Average coverage(se)	Coverage below/above	Average length
1	n	Empirical	95%	2.5%/2.5%	3.65
	25	BJ	93.11(0.05)	0.38/6.51	3.83(1.04)
		CB	89.91(0.11)	5.72/4.37	3.75(1.23)
		PRR	93.17(0.08)	2.62/4.21	3.89(1.26)
	50	BJ	93.90(0.04)	0.24/5.87	3.88(0.72)
		CB	91.13(0.10)	5.10/3.77	3.68(0.94)
		PRR	94.01(0.07)	2.51/3.48	3.81(0.88)
	100	BJ	94.41(0.02)	0.03/5.56	3.89(0.53)
		CB	92.74(0.07)	4.17/3.10	3.70(0.64)
		PRR	94.79(0.05)	2.21/3.00	3.80(0.68)
3	n	Empirical	95%	2.5%/2.5%	19.05
	25	BJ	96.55(0.04)	0.30/3.15	27.11(7.91)
		CB	89.25(0.10)	5.66/5.09	18.25(5.61)
		PRR	92.42(0.08)	2.98/4.60	19.77(6.23)
	50	BJ	97.51(0.02)	0.03/2.45	27.64(5.37)
		CB	92.01(0.08)	4.50/3.49	19.21(3.99)
		PRR	93.96(0.06)	2.72/3.32	19.99(4.10)
	100	BJ	97.83(0.01)	0.00/2.16	27.76(3.85)
		CB	93.47(0.05)	3.48/3.04	19.09(2.80)
		PRR	94.70(0.04)	2.36/2.95	19.64(2.99)

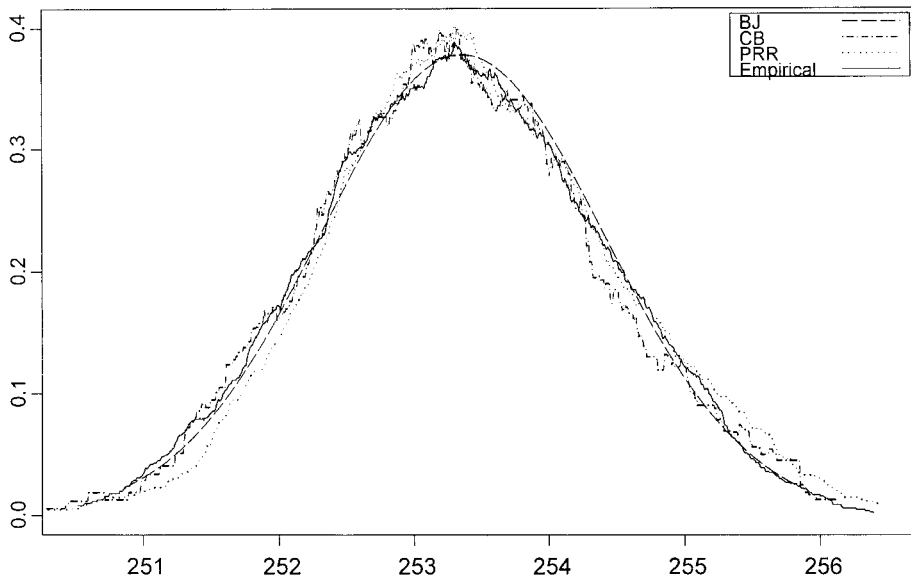
^aQuantities in parentheses are standard deviations.

Fig. 2. One-step ahead prediction densities of one series of size 240 generated by model (13) with Gaussian innovations.

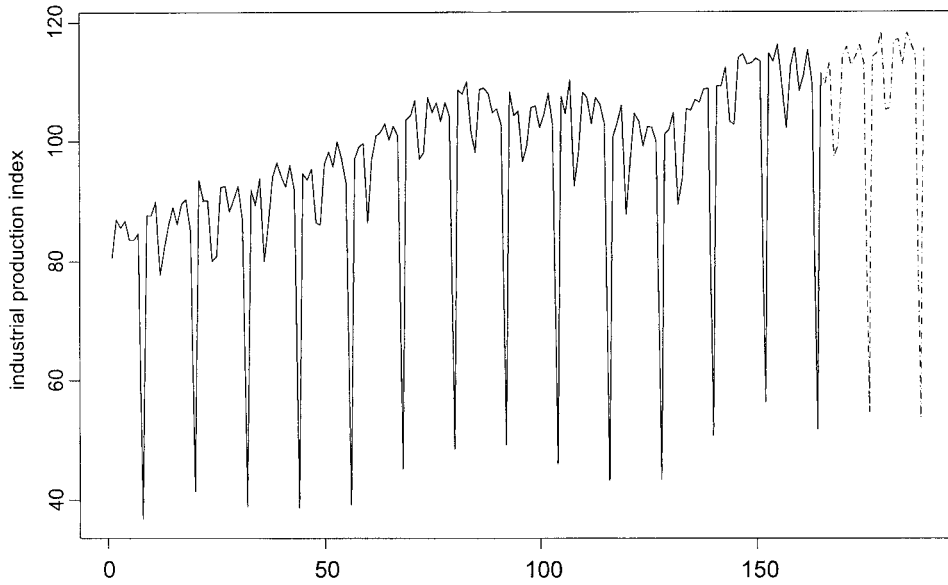


Fig. 3. Italian Industrial Production Index observed monthly from January 1983 to September 1998. Continuous line corresponds to estimation period and dashed line to prediction period.

The standard deviations in parentheses have been calculated using the asymptotic distribution. Residuals from model (19) have skewness coefficient -0.099 and excess kurtosis of -0.16 , so the Gaussianity hypothesis is not rejected at any usual level. Since the distribution of the residuals is not far from normality and the sample size is large, the intervals constructed using the BJ and PRR approaches are very similar. In Fig. 4, where 95% prediction intervals for \hat{y}_{T+k} , $k = 1, \dots, 24$, are plotted together with the actual observations and the linear point predictions, it can be seen that BJ and PRR intervals essentially coincide.

Next, we analyze the levels of the luteinizing hormone measured in a healthy woman every 10 min during 8 h. The data set is studied by Efron and Tibshirani (1993) and has been plotted in Fig. 5. The first 40 observations have been used to estimate the model to obtain

$$y_t = 1.19 + 0.48y_{t-1} + a_t.$$

(0.36) (0.16)

A kernel estimate of the residuals density

together with the normal density appear in Fig. 6. The empirical distribution of the residuals has a long tail to the right. The skewness coefficient is 0.83 and the excess kurtosis is 0.20 , both significantly different from the values under normality. We implement our procedure to construct the prediction density of the luteinizing hormone k steps ahead for $k = 1, \dots, 8$. The estimated densities for $k = 1$ and 3 appear in Fig. 7, with the asymmetry observed in the residuals distribution; see Fig. 6. Finally, from these densities we construct prediction intervals. Fig. 8 provides the point linear prediction, the observed levels of hormone and 80% and 95% prediction intervals constructed using Box–Jenkins and bootstrap procedures. It is clear the improvement in constructing prediction intervals using the PRR procedure over standard intervals. The 80% Box–Jenkins prediction intervals contains 3 out of 8 observations while the PRR intervals are able to cope with the asymmetry in the error distribution and include 5 observations without increasing the length of

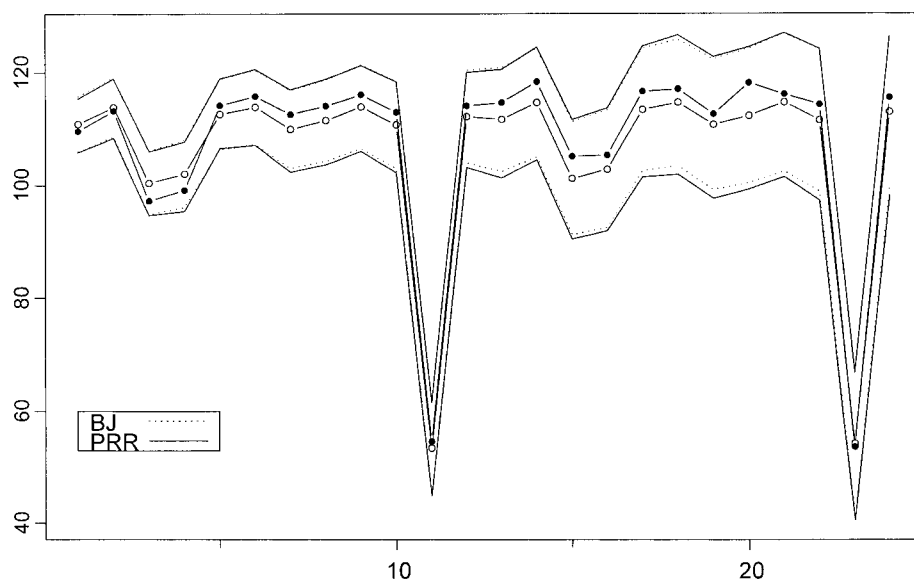


Fig. 4. Real observations of IPI (●) together with point linear predictions (○). 95% prediction intervals constructed by Box-Jenkins and bootstrap procedures.

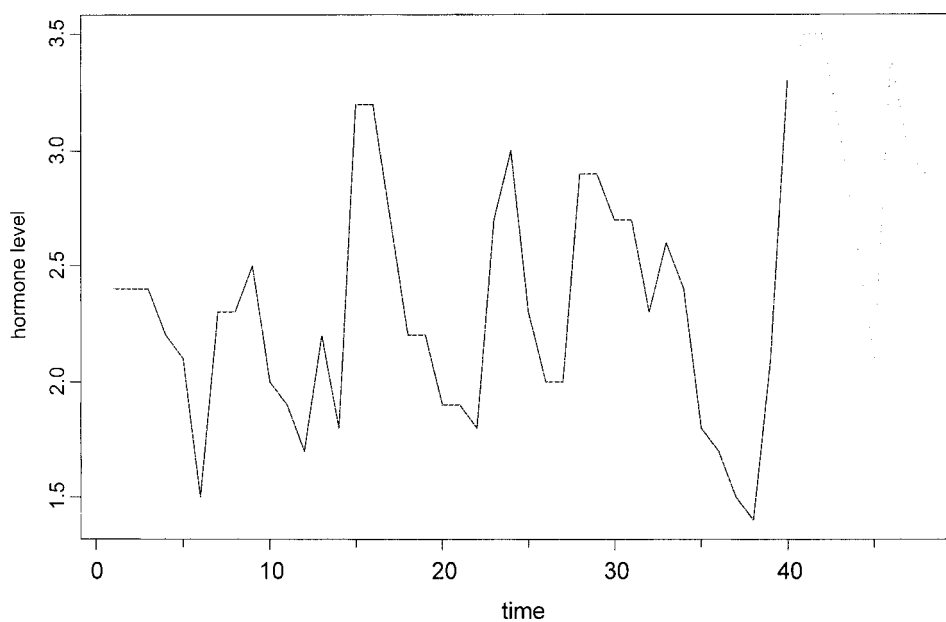


Fig. 5. Observations of the luteinizing hormone measured in a healthy woman every minute during 8 h. Continuous line corresponds to estimation period and dashed line to prediction period.

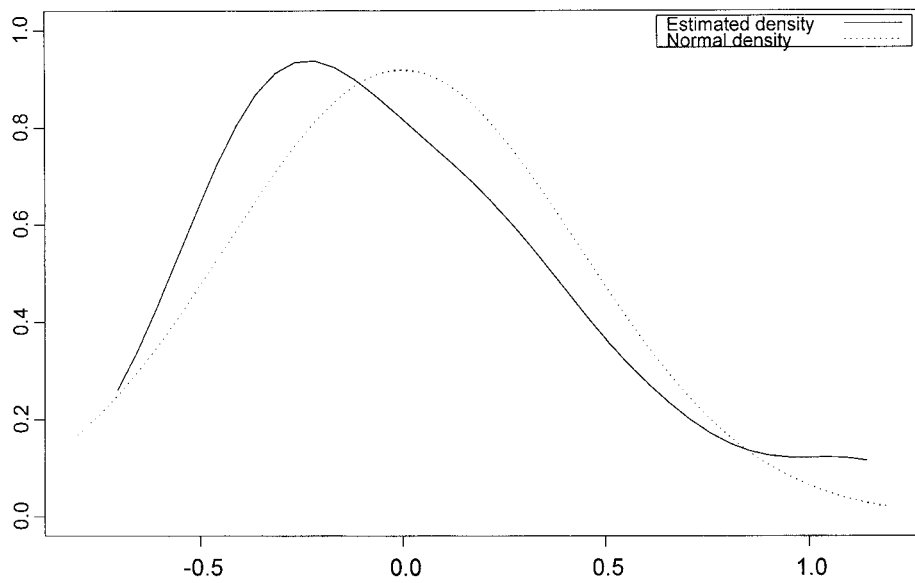


Fig. 6. Estimated density of residuals from AR(1) model for the luteinizing hormone and normal density.

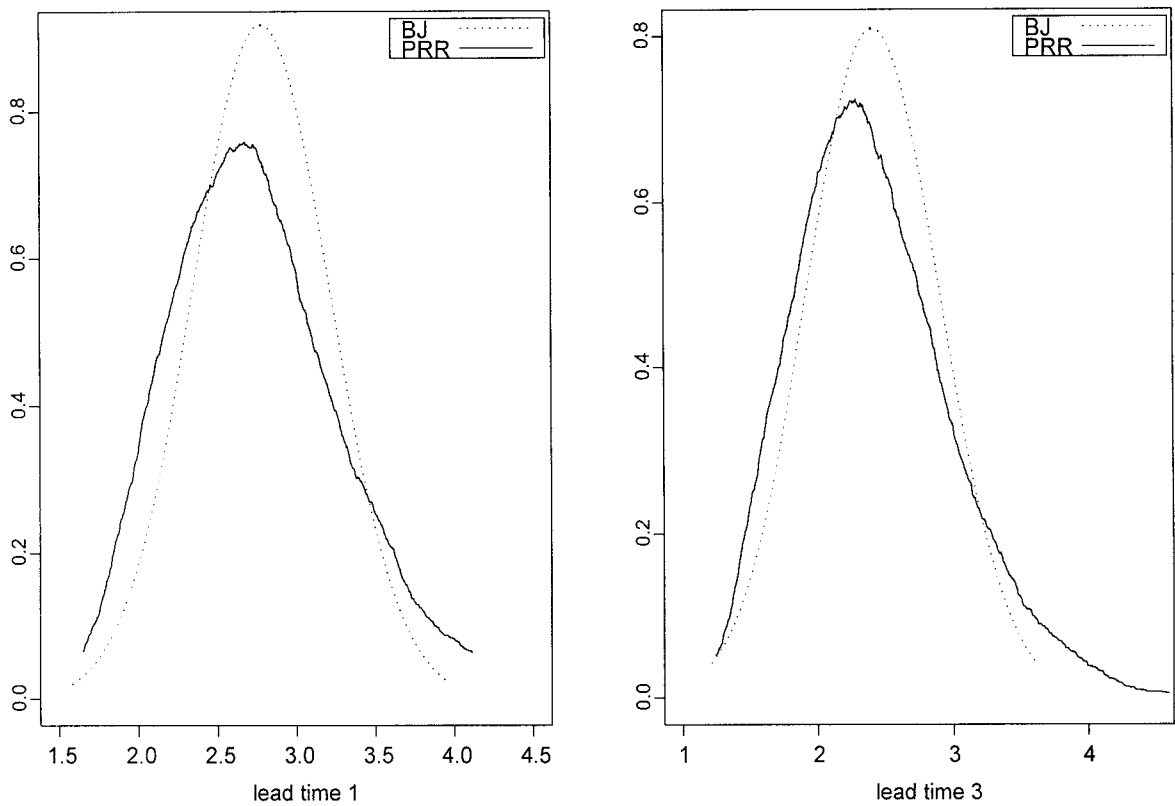


Fig. 7. Densities of one and three steps ahead predictions of the luteinizing hormone constructed by BJ and PRR procedures.

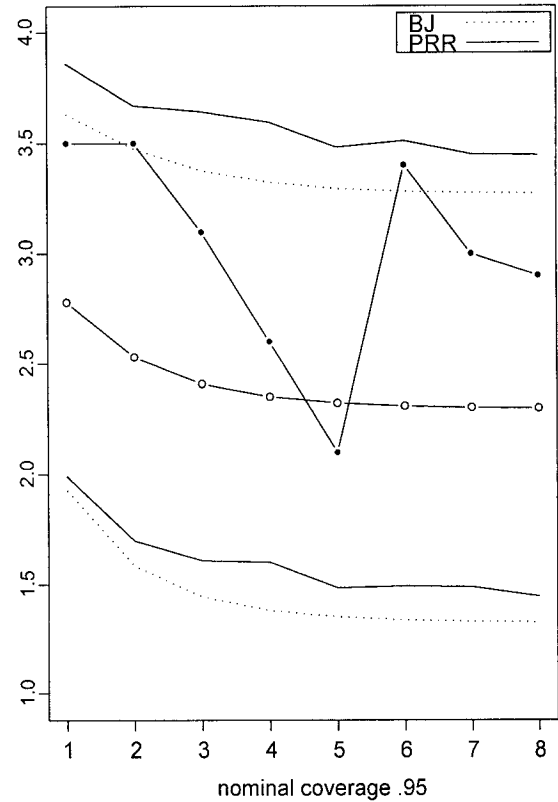
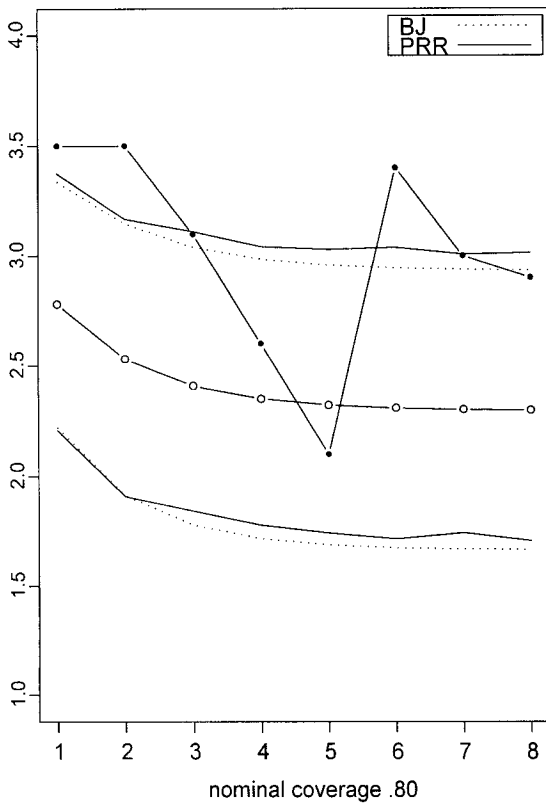


Fig. 8. Observations of luteinizing hormone (●) and point linear predictions (○). 80% and 95% intervals constructed by BJ and PRR procedures.

the intervals. Even when looking at the 95% prediction intervals, BJ intervals leave out 2 observations while PRR intervals do not leave out any observation. We have also computed bootstrap prediction intervals conditional on parameter estimates (CB). However, although the sample size is small, CB intervals are hardly distinguishable from PRR intervals and, consequently, we have not plotted them in Fig. 8. Therefore, it seems that for the values of the luteinizing hormone analyzed in this paper, the difference between BJ and PRR intervals is due to non-normality of the errors and not to parameter estimation. Efron and Tibshirani (1993) give the bootstrap distribution of the

OLS autoregressive parameter estimates with observations centered at the sample mean and using all 48 observations; the bootstrap standard error for $\hat{\phi}$ based on 200 bootstrap replicates is 0.12. The standard deviation of $\hat{\phi}$ is rather small with respect to the standard deviation of the errors (0.43) and this could explain why the parameter variability does not affect the shape of prediction intervals. This example shows how for small sample sizes and non-normal innovation distributions it could be worth considering bootstrap prediction intervals in order to improve the prediction performance of ARIMA models.

Since the error distribution of the luteinizing

hormone is not Gaussian, we have also estimated the parameters of the AR(1) model by LAD with the following results:

$$y_t = \underset{(0.37)}{0.73} + \underset{(0.17)}{0.68}y_{t-1} + a_t.$$

The standard deviations in parentheses have been calculated using the suggestion by Bassett and Koenker (1976). The point linear predictions of the luteinizing hormone provided by the OLS and LAD estimators have been plotted in Fig. 9. It can be seen that LAD predictions are systematically larger than OLS predictions and usually closer to the observed values. The MSE of the OLS predictions is 0.51 while for the

LAD estimator the prediction MSE is reduced to 0.38. Fig. 9 also represents the PRR 80% and 95% bootstrap intervals constructed using the OLS and LAD estimators. The LAD intervals adapt better to the asymmetry of the error distribution than the OLS intervals. Remember that the 80% PRR intervals constructed with OLS estimates leave out 3 observations when they were supposed to leave approximately one out. When PRR intervals are constructed using LAD estimates, they leave out only one observation. Consequently, it seems, as expected, that using parameter estimators more appropriate to the innovations distribution improves the performance of PRR prediction intervals.

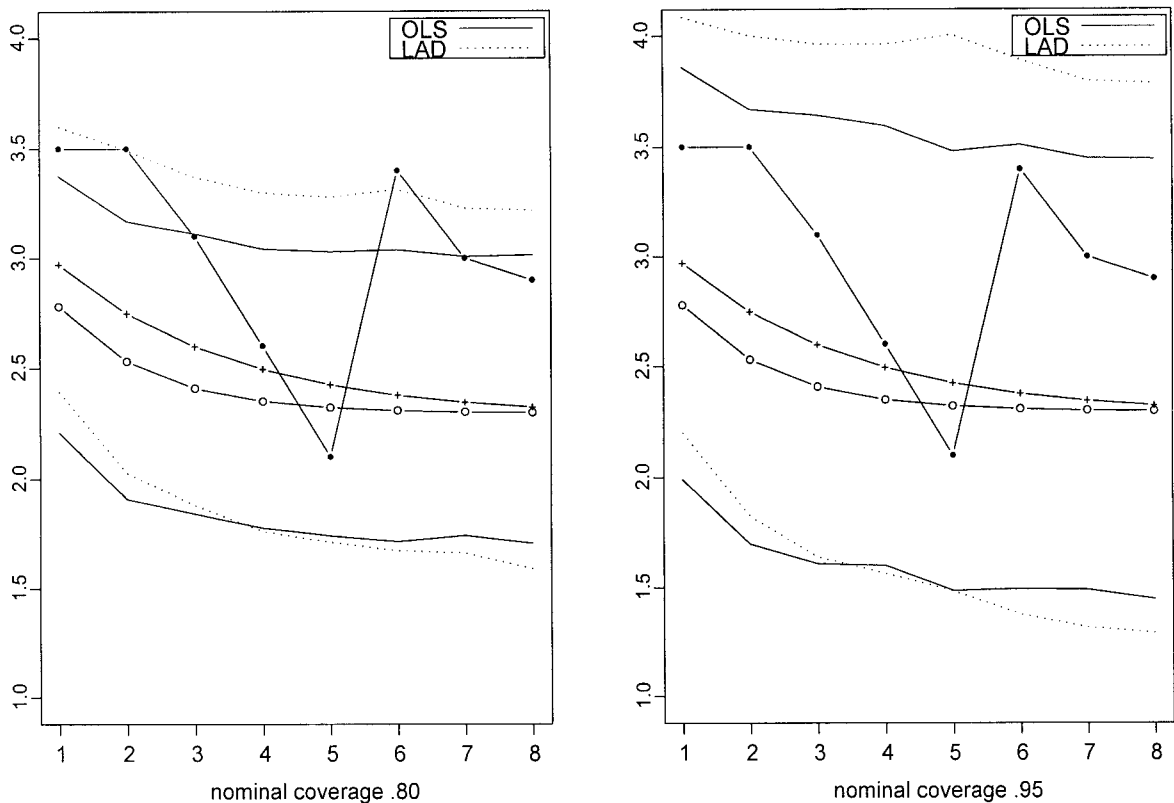


Fig. 9. Observations of luteinizing hormone (●) and point linear predictions obtained using OLS (○) and LAD (+). 80% and 95% intervals constructed using OLS and LAD.

5. Conclusions

This paper focuses on the effects of parameter estimation on the shape of prediction densities for seasonal ARIMA models. Prediction intervals can be obtained using bootstrap procedures which do not assume any distribution for the errors and can incorporate the variability due to parameter estimation. In particular, we consider the bootstrap technique proposed by Pascual et al. (1998) for ARIMA models extending it to models with seasonal components. By means of Monte Carlo experiments, we have first studied how coverage and length of prediction intervals are affected by not taking into account the variability due to parameter estimation. We show that the average coverage of the intervals is closer to the nominal value when intervals are constructed incorporating parameter uncertainty. As expected, since we are considering consistent estimators, the effects of parameter estimation are particularly important for small sample sizes. Furthermore, these effects are more important when the error distribution is not Gaussian. We also analyze the effect of the estimation method on the shape of prediction densities. In particular, we compare prediction densities constructed when the parameters of $ARI(p, d)$ models are estimated by OLS and by LAD. We show how, when the error distribution is not Gaussian, the average coverage and length of intervals based on LAD estimates are closer to nominal values than those based on OLS estimates. Since both estimates are consistent, this effect is less evident as the sample size grows. It is remarkable how bootstrap prediction intervals adapt to the asymmetry of the problem providing asymmetric prediction intervals improving on the necessarily symmetric BJ prediction intervals (see, e.g., Fig. 8).

Finally, the performance of the PRR technique is illustrated with two empirical examples. First, we estimate prediction densities for a

monthly series of the Italian IPI. Since the sample size is rather big (165 observations) and the innovation distribution is not far from normality, the prediction intervals obtained by BJ and PRR procedures are very similar. However, BJ and PRR prediction intervals constructed for the levels of a luteinizing hormone differ significantly

Several questions remain open for further research using resampling techniques; for example, the effect of the uncertainty on the specification of the model over prediction densities. This question has been addressed for autoregressions and using a different bootstrap strategy by Masarotto (1990) and Grigoletto (1998). However, they center the prediction intervals at a linear combination of past observations and this strategy may not be adequate when the distribution of the innovations is not Gaussian.

Acknowledgements

The authors are very grateful to the referees for comments that helped to improve the paper. Also, we are grateful to Lory A. Thombs for her comments, to Eva Senra for providing the Italian IPI data and to Regina Kaiser for helping us to clean the series of outliers. Financial support was provided by the European Union project ERBCHRXCT 940514 and by projects CICYT PB95-0299, DGICYT PB96-0111 from the Spanish Government and Cátedra de Calidad BBV.

References

- Bassett, G., & Koenker, R. (1976). Asymptotic theory of least absolute error regression. *Journal of the American Statistical Association* 73, 618–622.
- Box, G. E. P., & Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco.

- Cao, R., Febrero-Bande, M., Gonzalez-Manteiga, W., Prada-Sanchez, J. M., & Garcia-Jurado, I. (1997). Saving computer time in constructing consistent bootstrap prediction intervals for autoregressive processes. *Communications in Statistics, Simulation and Computation* 26, 961–978.
- Chatfield C. (1993). Calculating interval forecasts. *Journal of Business and Economic Statistics* 11, 121–135.
- Diggle, P. (1990). *Time Series. A Biostatistical Introduction*, Clarendon Press, Oxford.
- Efron, B., & Tibshirani, R. (1993). *An Introduction to the Bootstrap*, Chapman and Hall, New York.
- Gómez, V., & Maravall, A., (1996). Programs TRAMO and SEATS: Instructions for the user. Documento de Trabajo no. 9628, Banco de España (Spain).
- Grigoletto, M. (1998). Bootstrap prediction intervals for autoregressions: some alternatives. *International Journal of Forecasting* 14, 447–456.
- Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*, Springer-Verlag, New York.
- Masarotto, G. (1990). Bootstrap prediction intervals for autoregressions. *International Journal of Forecasting* 6, 229–239.
- Pascual, L., Romo, J., & Ruiz, E. (1998). Bootstrap predictive inference for ARIMA processes. Working Paper 98-86, Universidad Carlos III de Madrid.
- Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (1986). *Numerical Recipes*, Cambridge University Press, Cambridge.
- Stine, R. A. (1987). Estimating properties of autoregressive forecasts. *Journal of the American Statistical Association* 82, 1072–1078.
- Thombs, L. A., & Schucany, W. R. (1990). Bootstrap prediction intervals for autoregression. *Journal of the American Statistical Association* 85, 486–492.
- Biographies:** Lorenzo PASCUAL is Ph.D. student at the Departamento de Estadística y Econometría of Universidad Carlos III de Madrid. He works on bootstrap methods for prediction in time series models.
- Juan ROMO received his Ph.D. in Mathematics from Texas A&M University. He has been Associate Professor both in Universidad Complutense and Universidad Carlos III de Madrid, where he is Professor of Statistics and Operations Research since 1996.
- Esther RUIZ received her Ph.D. in Statistics from the London School of Economics in 1992. She is Associate Professor of Economics at Universidad Carlos III de Madrid since 1995.